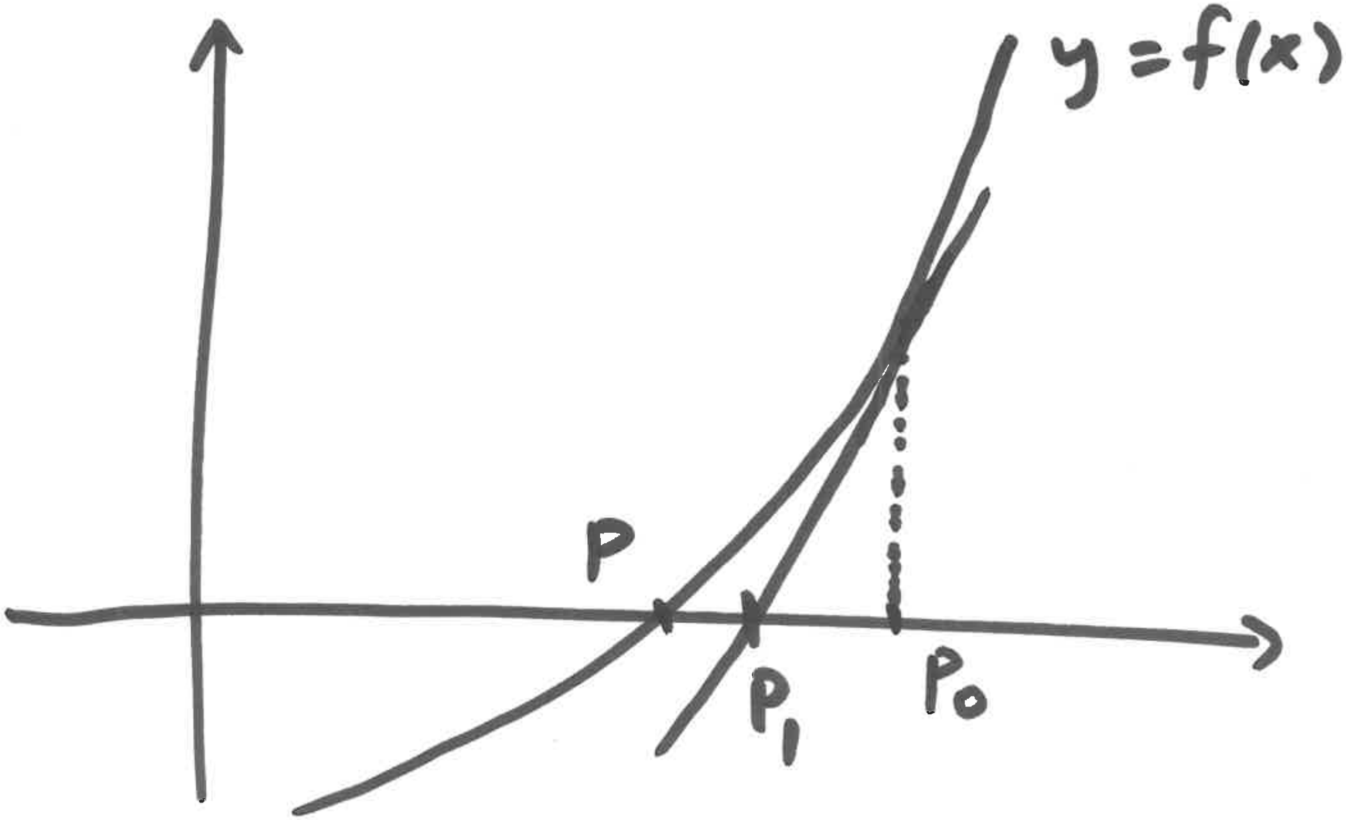
2.3 Newton's method

Find the root p of f



Suppose $|p_0 - p|$ is small, then near p_0 $f(p)$ is approximated by

$$0 = f(p) \approx f(p_0) + f'(p_0)\cdot(p - p_0)$$

Setting RHS $= 0$, we get

$$p \approx p_1 = p_0 - \frac{f(p_0)}{f'(p_0)}$$

<u>Newton's method</u>

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})} \qquad n \geq 1$$

- One expects that Newton's method is only good when $|p_0 - p|$ is small

- Newton's method is a special case of fixed point iteration :

$$p_n = g(p_{n-1}) \quad , \quad g(x) = x - \frac{f(x)}{f'(x)} .$$

$$g'(x) = 1 - \frac{f'(x) \cdot f'(x) - f(x) \cdot f''(x)}{(f'(x))^2}$$

$$g'(p) = 1 - \frac{(f'(p))^2 - 0}{(f'(p))^2} = 0$$

Thm Let $g \in C^2$. Suppose p is a fixed point of g with $g'(p) = 0$

Then $\exists\, \delta > 0$ s.t. the fixed point iteration starting inside $[p-\delta, p+\delta]$ converges to p, with

$$|p_{n+1} - p| \le C\,|p_n - p|^2 \quad \forall\, n \ge 0$$

for some $C > 0$.

"quadratic convergence"

pf) ① Convergence:

by continuity of g', $\exists\ \delta > 0$ s.t.

$$|g'(x) - g'(p)| \le \frac{1}{2} \quad \forall\ |x-p| \le \delta$$

$g'(p)=0$

$$\Rightarrow \quad |g'(x)| \le \frac{1}{2} \qquad x \in [p-\delta, p+\delta]$$

$$|g(x) - \underbrace{g(p)}_{\overset{\shortparallel}{p}}| = |g'(\xi)| \cdot |x-p|$$

$$\le \frac{1}{2}\delta$$

$$|g(x) - p| \le \frac{1}{2}\delta \le \delta \qquad g(x) \in [p-\delta, p+\delta].$$

By f. p. thm, $p_n \longrightarrow p$

② Quadratic convergence rate:

$$p_{n+1} = g(p_n)$$

$$p_{n+1} - p = g(p_n) - g(p)$$

$$= \cancel{g(p)} + \underbrace{\cancel{g'(p)}}_{=0} \cancel{(p_n - p)} + \frac{g''(\xi)}{2} \cdot (p_n - p)^2 \cancel{- g(p)}$$

$$= \frac{g''(\xi)}{2} (p_n - p)^2$$

g'' is continuous $\Rightarrow \sup_{x \in [p-\delta, p+\delta]} |g''(x)|$ is finite.

$$|p_{n+1} - p| \le C\,|p_n - p|^2$$

$$C = \frac{1}{2} \sup_{x \in [p-\delta,\, p+\delta]} |g''(x)|$$

- - - - - - - -

Advantages of fixed point iteration:

- Can be generalized to multi-D.
- Newton's can reach quadratic convergence.

Disadvantages:

- The contraction property is hard to check.

(for Newton's, convergence is only guaranteed near p: "conditional stability")

- For Newton's, $f'(x)$ is needed. This can be expensive in multi-D:

$$\vec{f}(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_n(\vec{x}) \end{pmatrix}$$

$$\nabla \vec{f}(\vec{x}) = \Big(\partial_{x_i} f_j(\vec{x}) \Big)_{n \times n}$$